

TEST-RETEST RELIABILITY OF ASSESSMENT OF ANAESTHETIST'S PERFORMANCE IN MANAGING SERIOUS ADVERSE EVENTS ON A SCREEN-BASED COMPUTER SIMULATOR

Jyotirmay Kirtania¹, Shreyasi Ray²

¹Associate Professor, Department of Anaesthesiology, ESI-PGIMSR and ESIC Medical College, Joka, Kolkata.

²Assistant Professor, Department of Anaesthesiology, ESI-PGIMSR and ESIC Medical College, Joka, Kolkata.

ABSTRACT

BACKGROUND

Serious Adverse Events (SAE) during anaesthesia procedure are uncommon, but potentially fatal. Traditional methods to assess the performance of anaesthetists in SAE management are largely subjective and unreliable. In this study, we assessed the test-retest reliability of the assessment of performance of anaesthetists in managing SAE during anaesthesia procedure by determining if different observations assessing the same clinical scenario on a computer screen based anaesthesia simulator would have agreement on the rating of the anaesthetist's performance.

MATERIALS AND METHODS

Three SAE scenarios on a computer screen based simulator were used for assessment. From a pilot study, the minimum number of trials beyond which there is little or no improvement of performance was determined for each SAE scenario. Six fresher anaesthesia postgraduates were enrolled. Each participant undertook ten trials for each SAE scenario and submitted the trial-logs in PDF format to the investigator. SAE detection latency, initial treatment latency period and SAE treatment completion period were collected from the simulator trial logs. Test-retest reliability (internal consistency) between trials was assessed with Pearson's Product Moment Correlation Coefficient. Scatter plot of improvement of latency period and treatment time versus the number of simulator trials were used to characterise the learning curves. Regression model fitting of the learning curve was used to find the best fit model for the learning curve.

RESULTS

There was significant test-retest reliability and internal consistency between the trials for each of the three SAE scenarios.

CONCLUSION

PC based simulated SAE scenarios are effective and reliable for the assessment of performance of fresher anaesthesia postgraduates.

KEYWORDS

Serious Adverse Events, Anaesthesia, PC Based Simulation, Test-Retest Reliability, Performance Assessment.

HOW TO CITE THIS ARTICLE: Kirtania J, Ray S. Test-retest reliability of assessment of anaesthetist's performance in managing serious adverse events on a screen-based computer simulator. J. Evolution Med. Dent. Sci. 2017;6(79):5575-5579, DOI: 10.14260/jemds/2017/1210

BACKGROUND

Serious adverse events (SAE) during anaesthesia procedure are uncommon. But SAE during anaesthesia can be rapidly fatal or can cause permanent disability to the patient unless promptly recognised and corrected by the anaesthesia caregiver.^[1]

Traditional anaesthesia training for SAE management largely involve the cognitive domain with little or no practice of the higher cognitive domain, psycho-motor domain and affective domain. This is due to the fact that it is ethically not permissible to expose patients to SAE under anaesthesia for the purpose of training, as also the uncommon occurrence of SAE during anaesthesia. Since SAE under anaesthesia are uncommon, anaesthetists usually "learn" SAE management throughout their career, sometimes risking patient's lives and consequent medico-legal actions.

Thus, the traditional process of learning to manage SAE during anaesthesia procedure is no longer tenable in the modern day anaesthesia practice.

The theory of the learning curve is based on the simple idea that the competence increases, and time required to perform a task decreases as the worker gains experience. High risk industries like aviation and space technology, constantly train and assess their personnel on simulators to improve and ensure the safety of their customers against the occurrence of uncommon, but serious adverse incidents involving loss of life or property. Although, simulator-based learning is not new in Anaesthesia, the learning curve of fresher anaesthesia-postgraduates in managing serious adverse events during anaesthesia procedure on a computer simulator was not studied till date.

A learning curve is a collection of data points (x_j , y_j), which describe how the performance (y_j) is related to training sample sizes (x_j), where $j = 1$ to m , m being the total number of instances. These learning curves can typically be divided into three sections: In the first section, the performance increases rapidly with an increase in the size of the training set; the second section is characterised by a turning point where the increase in performance is less rapid and a final section where the classifier has reached its efficiency threshold, i.e. no (or only marginal) improvement

'Financial or Other Competing Interest': None.
Submission 09-09-2017, Peer Review 23-09-2017,
Acceptance 25-09-2017, Published 30-09-2017.

Corresponding Author:

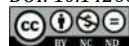
Dr. Shreyasi Ray,

B 111, Survey Park,

Santoshpur, Kolkata, West Bengal, India.

E-mail: rayshreyasi@rediffmail.com

DOI: 10.14260/jemds/2017/1210



in performance is observed with increasing training set size.^[2]

In this study, we assessed the test-retest reliability of the anaesthetists' performances on a computer screen based SAE simulator by determining if different observations assessing the same SAE scenario would have agreement on the rating of the anaesthetist's performances. This study also aims to characterise the learning curve of fresher anaesthesia-postgraduates in managing serious adverse events during anaesthesia procedure on a simulator.

MATERIALS AND METHODS

The study was conducted between June and September 2016, after seeking approval from the Institute Ethical Committee. The study design was prospective observational - non-randomised allocation, same participant act as control and test with repetitive measures.

A survey among four anaesthesia faculty members was carried out to prepare a validated list of the minimum number (N_s) of SAE scenarios for the simulator training. It was found that $N_s = 3$ (three). These three SAE scenarios were: (i) Anaphylaxis, (ii) ST-T changes and (iii) Obstructed airway. Using this validated list of the minimum number ($N_s = 3$) of SAE scenarios, a pilot study was conducted on two senior anaesthetists (> 6 years post MD experience), to determine the number of trials beyond which there is little or no improvement of performance (time to detect SAE, initiate treatment and complete treatment for each scenario). Using this pilot study data, the minimum number of trials (N_t) to be tested for each SAE scenario was determined to be $N_t = 10$. Thus, the sample size (number of simulator trials or simulator trial logs) to detect differences between performances on repeated test trials was at least $= (N_s \times N_t \text{ or } 3 \times 10) = 30$, for each participant. Six (6) fresher anaesthesia-postgraduates were enrolled in the study. Thus, the sample size (the number of simulator trials) allocated to each of the six participants, $m = (N_s \times N_t) = 30$, and the total sample size (total number of simulator trials) was $6 \times 30 = 180$.

Stratified Random Sampling was done for the participants: A list of 60 (sixty) senior residents in Anaesthesiology was prepared from the data available at <https://mciindia.org/BrowseBYDept.aspx> (accessed June 2016; database currently moved to <https://old.mciindia.org/BrowseBYDept.aspx>) under the "Information Desk" link on the Medical Council of India (MCI) website. Only fresher postgraduate senior residents (less than 1 year post-MD experience) in Anaesthesiology fulfilling inclusion criteria (below) were approached for participation and nine willing participants fulfilled the inclusion criteria. The participants were contacted through their Institute's contact data available on the public domain database of Institute's website they work. Of the nine willing participants fulfilling the inclusion criteria, six participants were finally enrolled in the study using a computer generated random number table.

Consecutive sampling was done for the simulator data: All the completed simulator trial logs (satisfying the inclusion criteria) were included in the study till the required sample size (training set size) is achieved.

Inclusion Criteria for Simulator Logs

1. Completed simulator logs for the validated SAE scenarios.

Inclusion Criteria for Participants

1. Fresher postgraduate senior residents in Anaesthesiology with less than 1 year post-MD experience.
2. Senior residents in Anaesthesiology, having passed MD in the first attempt.
3. Participants with basic knowledge of Windows based PC systems.

Exclusion Criteria for Simulator Logs

1. Incomplete simulator logs.
2. Simulator logs from non-validated SAE scenarios.

Exclusion Criteria for Participants

1. Postgraduate senior residents with more than 1 year post-MD experience.
2. Non-MD senior residents (e.g. with DA, DNB).
3. Senior residents having passed MD on more than one attempt.
4. Senior residents with additional training in Anaesthesiology (e.g. DA, DNB, Junior Residency) prior to MD.
5. Participants without basic knowledge of Windows based PC systems.

The Anesoft Anaesthesia Simulator 6™ (Windows™ PC based simulator) was used for the study.

In the learner sensitisation phase of the study, the study participants were instructed to use the simulator by the principal investigator (PI) and also by self-directed learning by participants using the "Consultant" module of the simulator. During each practice trial, one SAE scenario (from the validated list) was manually selected by the participant. Each participant undertook 'm = 10' number of training sessions for each of the three SAE scenarios over a period of six weeks. At the end of each simulator trial, the participant saved the PDF log of the session with a unique file name and emailed it to the PI. The participant also viewed the simulator log to get appropriate feedback regarding his/her performance in that simulator session.

The following Data were to be Collected for each Participant

Demographic Data: (Data Tool: Data from participant-Enrolment-Form)

- Age.
- Sex.
- Date of passing MD Anaesthesiology.
- Prior training on anaesthesia simulator (Yes/No).
- Basic knowledge of windows operating system and email (Yes/No).

Simulator Data (Data Tool: Time and event log of each simulator session generated by the simulator and emailed by the participant to the principal investigator).

The following Primary Data from the Simulator were Collected for each Practice Session from Assessment of the Simulator Session Log

- Practice session start time.
- SAE start time.
- SAE detection time.
- Initial treatment start time.
- Treatment end time.
- All the recommended treatment steps completed or not.

The following Derived Data was calculated from the Simulator Log for each Simulator Session for Analysis

- The response time of the study participant to detect SAE ("SAE detection latency"). It was defined as the interval between the occurrence of the SAE and the detection by the study participant during each practice session.
- The response time of the study participant for initial action to treat the SAE ("initial treatment latency"). It was defined as the interval between the occurrence of the SAE and its initial corrective action taken by the study participant during each practice session.
- Whether the participant correctly performed all the steps of the standard treatment for the SAE (yes/no).
- The time taken by the study participant to perform all the corrective measures for the SAE (if correctly completed at all). "SAE treatment completion time" was defined as the interval between the start of the initial step of treatment and the completion of the last step of treatment performed by the participant.
- Learning curve of participants (scatter plot) for each scenario was constructed based on improvement over successive simulator trials using the "SAE detection latency," "initial treatment latency" and "SAE treatment completion time."

Statistical Methods

Test-retest reliability (internal consistency) between trials was assessed with Pearson's product moment correlation coefficient alpha. P value of < 0.05 was considered statistically significant. Scatter plot of number of practice sessions versus participant performance (improvement of latency period and treatment time) was used to characterise the learning curve for each SAE scenario. Regression analysis was used for model fitting (to characterise the best fitting distribution model) of the scatter plots of the learning curves.

Statistical Analysis was done exclusively with "LibreOffice Calc" version 5.1.4.2 (2016) and "R" version 3.3.1 (2016) with package "Rcmdr" version 2.2.5 (2016), all of which are free and open-source software.

RESULTS

The descriptive statistics of demographic profile and professional experience of the experts (pilot study participants) and the fresher postgraduates (study participants) are described in Table 1. All the experts and fresher postgraduate participants completed the stipulated number of simulator trials for each SAE scenarios. Every expert and participant completed all the recommended steps for treatment of the SAEs in each of the simulator trials.

Test-retest reliability (internal consistency) of performance data (SAE detection latency, initial treatment latency and SAE treatment completion time) between trials was assessed with Pearson's product moment correlation coefficient alpha. Closer the alpha value towards unity (1.0), greater the test-retest reliability of the performance data. Test-retest reliability measures of performance of experts in the pilot study were: Alpha reliability = 0.7933; Standardised alpha = 0.6752. Test-retest reliability measures of performance of fresher postgraduates in the study were: Alpha reliability = 0.7957; Standardised alpha = 0.6681. So these simulator trials were reliable assessment tools to assess the learning (improvement of performance) of fresher postgraduates, while they manage simulated SAE during anaesthesia procedures.

Scatter plots of various performance measurement (time) of experts versus trial numbers for each SAE scenarios are depicted in Figure 1. It is seen that the learning curve (scatter plot) of experts for each SAE scenario reached a plateau within ten successive trials on the simulator. Scatter plots of various performance measurement (time) of fresher postgraduates versus trial numbers for each SAE scenarios are depicted in Figure 2. It is seen that the learning curve (scatter plot) of fresher postgraduates for each SAE scenario also reached a plateau within ten successive trials on the simulator.

Regression analysis (Table 2, Table 3) shows that all the learning curves (scatter plots of performance measurement versus trial number) approximately fit the "power-law distribution model," rather than the linear or the logarithmic models.

	Experts	Fresher Postgraduates
Number of participants	2	6
Male: Female	1:1	3:3
Median age in years (Min., Max.)	53.5 (52, 55)	30 (29, 32)
Median clinical experience (post MD) in years (Min., Max.)	23.5 (22, 25)	0.67 (0.67, 0.92)

Table 1. Demographic Characteristics of Pilot Study Participants (Experts) and Study Participants (Fresher Postgraduates)

Type of Scatter Plot	Regression Model	Standard Error of the Regression (S)*	Slope (α)	Intercept (B)
SAE detection latency versus trial number	Linear	154.7	- 11.0	371.0
	Logarithmic	154.2	- 48.5	383.7
	Power	0.551*	- 0.149	337.4
Initial treatment latency versus trial number	Linear	245.5	- 15.9	535.2
	Logarithmic	244.9	- 70.0	553.7
	Power	0.675*	- 0.149	461.2
SAE treatment completion time versus trial number	Linear	380.4	- 39.5	1331.3
	Logarithmic	377.9	- 174.3	1377.3
	Power	0.405*	- 0.150	1294.3

Table 2. Overall Learning Curve (Performance Improvement) Characteristics of Experts during the Pilot Study for the Three SAE Scenarios

*Lesser the value of the standard error of the regression (S) of the data, better the data fits with that regression model. In this case, the best fit model is the Power Model, $T = B * n^{-\alpha}$

Type of Scatter Plot	Regression Model	Standard Error of the Regression (S)*	Slope (α)	Intercept (B)
SAE detection latency versus trial number	Linear	149.8	- 12.2	405.2
	Logarithmic	149.2	- 54.1	419.6
	Power	0.499*	- 0.157	380.4
Initial treatment	Linear	248.5	- 17.5	565.2
	Logarithmic	247.8	- 76.8	602.6

latency versus trial number	Power	0.633*	- 0.151	512.5
SAE treatment completion time versus trial number	Linear	374.2	- 43.1	1450.2
	Logarithmic	371.2	- 190.1	1500.4
	Power	0.357*	- 0.150	1429.4

Table 3. Overall Learning Curve (Performance Improvement) Characteristics of Fresher Postgraduates during the Study for the three SAE Scenarios

*Lesser the value of the standard error of the regression (S) of the data, better the data fits with that regression model. In this case, the best fit model is the Power Model, $T = B * n^{-\alpha}$

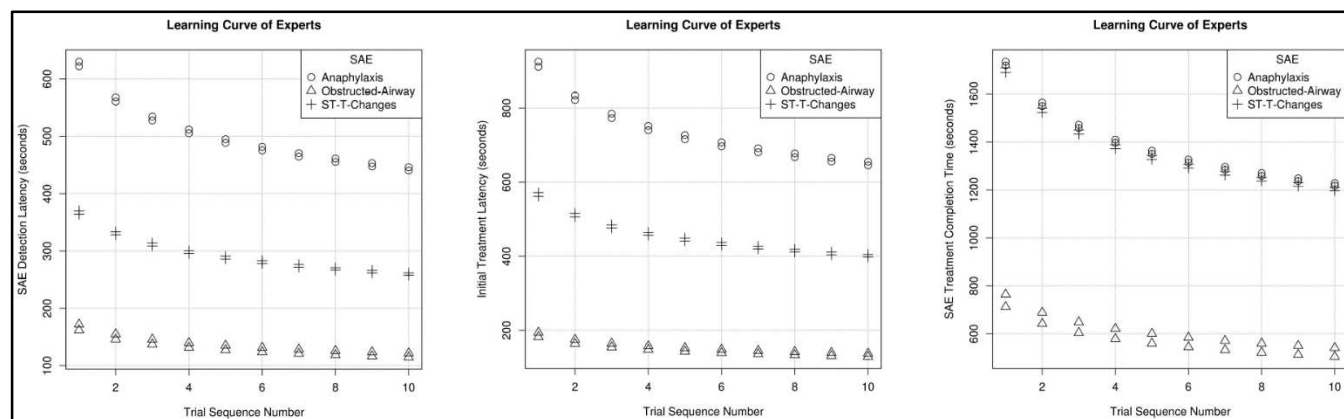


Figure 1. Scatter Plot showing Learning Curve of Experts

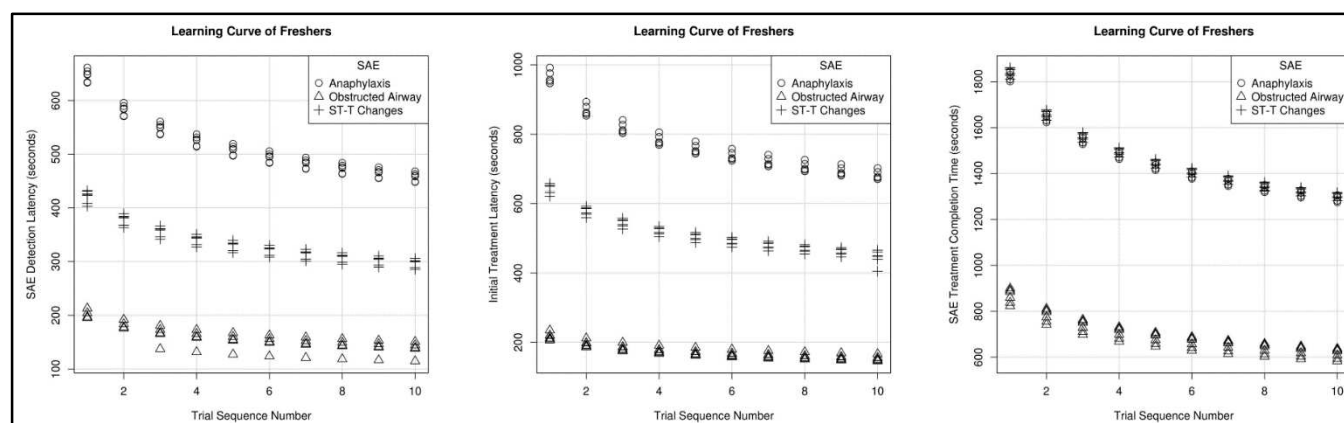


Figure 2. Scatter Plot showing Learning Curve of Fresher Postgraduates

DISCUSSION

The slope (steepness) of the learning curve for experts and fresher postgraduate are similar (Figure 1 and Figure 2). The intercepts (initial performance measures) are different (Table 2 and Table 3) for experts and fresher postgraduates. However, this study was not designed to detect quantitative differences in the performance measure of experts and fresher postgraduates.

In regression analysis, lesser the value of the standard error of the regression (S) of the data in a regression model, better the data fits with that distribution model. The coefficient of determination (R^2) is not a valid indicator of model fitting, particularly when the data follow non-linear model as in our study.^[3] So R^2 was not used for the model fitting. In our study, the learning curve approximately fits

with the power-law equation, $T = B * n^{-\alpha}$, where 'T' is the performance measurement (Time) and 'n' is the number of trials. The constant 'B' represent the performance measurement at the beginning of the trial (i.e. $n = 1$). The exponent (Power) ' α ' represent the 'slope' or 'steepness' of the learning curve (Scatter plot). Unlike an exponential curve, whose linear slope is a constant multiple of 'T', the slope of this "power-law" distribution curve varies with the number (n) of the trials. In this case, at higher values of 'n', the curve flattens out (plateau of the learning curve).

Traditional methods (written examination, oral case presentation and direct observation of procedural skills) to assess the performance of anaesthetists in SAE management are largely subjective and non-reproducible. The ability to measure actual performance, vigilance, interpretation of data

in real time and formulation and implementation of a management plan is not readily demonstrable by these traditional methods. Our study is novel, as it has proven that a computer screen based simulator can assess the performance of anaesthetists with a valid test-retest reliability.

The long-term impact of this simulator training on patient safety during anaesthesia care cannot be assessed by performance of anaesthetists on simulators. Hays et al (1992) conducted a meta-analysis of flight simulation research to identify important characteristics associated with the effectiveness of simulator training of military aircraft pilots. The major finding was that the use of simulators combined with aircraft training consistently produced improvements in training compared to aircraft training only.^[4]

However, in the field of anaesthesia, no study has been undertaken which investigated the efficacy of simulator training to improve the management of real life SAE scenarios. Again, ethical issues and the rarity of SAE under anaesthesia may prevent such studies to be undertaken. If widespread simulator training and assessment of SAE management are introduced for anaesthetists, the impact can be judged by studying the trend (over years or decades) in closed claim registries and medico-legal suits in anaesthesia practice.

CONCLUSION

The performance data (learning curve) of fresher anaesthesia postgraduates managing computer simulated SAE scenarios approximately follow a “power-law distribution” curve. For each SAE scenario, a performance plateau was observed within ten practice sessions. This simulator trial method has valid test-retest reliability for each of the three SAE scenarios. Hence, this method is valid and reliable for performance assessment of anaesthetists managing simulated SAE under

anaesthesia on a PC based simulator. This method of performance assessment should be incorporated into postgraduate medical curriculum in Anaesthesiology in addition to the traditional methods of assessment of practical skills like oral case presentation, viva-voce, direct observation of procedural skills (DOPS), and objective structured clinical examination (OSCE). For a given SAE scenario during anaesthesia procedure, ten simulation trials are adequate and reliable for the learning and assessment of learning. Long term retention of simulator acquired skills and efficacy of simulator training in improving patient safety need to be assessed by further studies.

ACKNOWLEDGEMENT

We are grateful to Professor (Dr.) Tripti Srivastava (Waghmare), Convener of MCI Nodal Centre for Medical Education, Datta Meghe Institute of Medical Sciences (DU), Wardha, Maharashtra, for her valuable guidance and support in this research project.

REFERENCES

- [1] Jenkins K, Baker AB. Consent and anaesthetic risk. *Anaesthesia* 2003;58(10):962–84.
- [2] Hopper AN, Jamison MH, Lewis WG. Learning curves in surgical practice. *Postgraduate Medical Journal* 2007;83(986):777-9.
- [3] Spiess AN, Neumeyer N. An evaluation of R2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a monte carlo approach. *BMC Pharmacology* 2010;10:6.
- [4] Hays RT, Jacobs JW, Prince C, et al. Flight simulator training effectiveness: a meta-analysis. *Military Psychology* 1992;4(2):63-74.